

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:52:21 ; Search time 170.72 seconds  
(without alignments)  
25.057 Million cell updates/sec

Title: US-09-331-631a-1\_COPY\_186\_248

Perfect score: 355  
Sequence: 1 KRDPQREYEDCRRRCEOE.....MNPQGGSGRYEGEEEOS 63

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR\_66:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	34.9	509	2	S08059 alpha-globulin typ
2	120	33.8	566	2	S22477 vicilin precursor
3	114	32.1	588	1	FWCNAB alpha-globulin B p
4	112	31.5	637	2	S35221 globulin Bgl prec
5	101	28.5	605	2	S06398 alpha-globulin typ
6	92	25.9	810	2	T44430 protein PVI00 (imp
7	86	24.2	1655	2	T13998 gene mastermind pr
8	84	23.7	966	2	S25365 CYC8 protein - yea
9	79.5	22.4	236	2	T01662 globulin-1 - maize
10	79	22.3	411	2	T29475 hypothetical prote
11	79	22.3	919	2	A39248 androgen receptor
12	78	22.0	1671	2	S71628 sensory transducti
13	77.5	21.8	613	2	S27770 hypothetical prote
14	77.5	21.8	910	2	A34721 androgen receptor
15	77.5	21.8	911	2	B34721 androgen receptor
16	77	21.7	646	2	D82493 conserved hypotet
17	75.5	21.3	330	2	T25169 hypothetical prote
18	75.5	21.3	834	2	T42702 hypothetical prote
19	75	21.1	1090	2	A41696 regulatory protein
20	74	20.8	1038	2	T02634 rep protein homolo
21	74	20.8	1898	1	A45973 trichosyalin - hum
22	73.5	20.7	185	2	S71512 hypothetical prote
23	73.5	20.7	582	2	B53234 vicilin-like stora
24	73	20.6	407	2	T02258 globulin-like stora
25	73	20.6	540	2	S21825 vicilin-like stora
26	73	20.6	562	2	T26242 hypothetical prote
27	73	20.6	568	2	T26243 hypothetical prote
28	73	20.6	573	2	A53234 globulin-15, GLB15
29	73	20.6	3190	2	T13828 CREB-binding prote

30	71.5	20.1	669	2	JC5662 hepatoma-derived g
31	71	20.0	139	2	A26892 Mopa box protein -
32	71	20.0	162	2	S49259 albumin 4 - easter
33	71	20.0	625	2	A34615 profilaggrin - rat
34	70.5	19.9	447	2	S52391 centrosomin B - mo
35	70.5	19.9	1344	2	T42637 hypothetical prote
36	69	19.4	133	2	I51960 Wtl1 zinc-finger ho
37	69	19.4	448	2	JC5046 Wtlms' tumor suppr
38	69	19.4	448	2	S33926 Wtlms' tumor prote
39	69	19.4	449	2	A38080 Wtlms' tumor suscep
40	69	19.4	449	2	A39692 Wtlms' tumor prote
41	69	19.4	1023	2	S12519 glutactin - fruit
42	68.5	19.3	679	2	B75262 conserved hypotet
43	68.5	19.3	758	2	S54522 hypothetical prote
44	68	19.2	1069	2	T00377 KIAA0642 protein -
45	68	19.2	1154	2	S69206 regulator protein

## ALIGNMENTS

RESULT 1  
S08059  
alpha-globulin type B precursor (tandem 1) - upland cotton (fragment)  
N:Alternate names: seed storage protein  
C:Species: Gossypium hirsutum (upland cotton)  
C>Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 30-Sep-1993  
C:Accession: S08059  
R:Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.  
Plant Mol. Biol. 9, 533-546, 1987  
A>Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.  
A:Reference number: S06398  
A:Accession: S08059  
A>Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-509 <CHL>  
C:Superfamily: glycinin

Query Match 34.9%; Score 124; DB 2; Length 509;  
Best Local Similarity 37.0%; Pred. No. 9.1e-06;  
Matches 34; Conservative 9; Mismatches 15; Indels 34; Gaps 5;

QY 3 DPQREYEDCRRRCEOEPRQHCQLRC-----RECO-----RHNRGSDMM--- 45  
|||:| ||:|:| ||| ||| ||| ||| |||  
Db 1 DPQRR-YEECQCEGRQGEERQCPCCQRCCLKRFEDQDSORPQECQCHQGEORPEK 59  
|||:| ||:|:| ||| ||| ||| ||| |||

QY 46 -----NPDQGGSGRYEGEEEOS 63  
|||:| ||:|:| ||| ||| ||| ||| |||

Db 60 KQCCVRCREKRYQENPMRG--EREFEAEFEET 89  
|||:| ||:|:| ||| ||| ||| ||| |||

RESULT 2  
S22477  
vicilin precursor - cacao  
C:Species: Theobroma cacao (cacao)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 17-Mar-1999  
C:Accession: S22477; S22478; S18105; S22050  
R:McHenry, L.; Fritz, P.J.  
Plant Mol. Biol. 18, 1173-1176, 1992  
A>Title: Comparison of the structure and nucleotide sequence of vicilin genes of coco  
A:Reference number: S22477; MUID:92288309  
A:Accession: S22477  
A:Molecule type: DNA  
A:Residues: 1-566 <MCH>  
A:Cross-references: EMBL:X62625  
A:Accession: S22478  
A:Molecule type: mRNA  
A:Residues: 1-452 <MC2>  
A:Cross-references: EMBL:X62626  
C:Genetics:  
A:introns: 211/1; 269/3; 296/3; 391/3; 502/1  
C:Superfamily: glycinin









